JA999 248
#2 RS

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されてる事項と同一であることを証明する。
This is to certify that the annexed is a true copy of the following application as filed h this Office.

出 願 年 月 日
ate of Application: ２０００年 ２月 ２日

願 番 号
plication Number: 特願２０００－０２５５９４

願 人
licant (s): インターナショナル・ビジネス・マシーンズ・コーポレーション

CERTIFIED COPY OF PRIORITY DOCUMENT

２０００年 ９月２２日

特許庁長官
Commissioner,
Patent Office 及川耕造

出証番号 出証特２０００－３０７７４０２

【書類名】　特許願
【整理番号】　JA999248
【あて先】　特許庁長官　殿
【国際特許分類】　G06F 13/00
【発明者】
【住所又は居所】　神奈川県大和市下鶴間１６２３番地１４　日本アイ・ビー・エム株式会社　東京基礎研究所内
【氏名】　依田　邦和
【発明者】
【住所又は居所】　神奈川県大和市下鶴間１６２３番地１４　日本アイ・ビー・エム株式会社　東京基礎研究所内
【氏名】　江藤　博明
【特許出願人】
【識別番号】　390009531
【住所又は居所】　アメリカ合衆国１０５０４、ニューヨーク州アーモンク（番地なし）
【氏名又は名称】　インターナショナル・ビジネス・マシーンズ・コーポレイション
【代理人】
【識別番号】　100086243
【弁理士】
【氏名又は名称】　坂口　博
【選任した代理人】
【識別番号】　100091568
【弁理士】
【氏名又は名称】　市位　嘉宏
【手数料の表示】
【予納台帳番号】　024154
【納付金額】　21,000円

【提出物件の目録】

【物件名】　明細書　1

【物件名】　図面　1

【物件名】　要約書　1

【包括委任状番号】　9706050

【包括委任状番号】　9704733

【プルーフの要否】　要

【書類名】　明細書

【発明の名称】　アクセス・チェーン追跡システム、ネットワーク・システム、方法、及び記録媒体

【特許請求の範囲】

【請求項１】　複数のコネクションから構成されるアクセス・チェーンを介して、ネットワーク上をパケットが通信されるシステムにおけるアクセス・チェーンの追跡方法であって、

第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較するステップと、

前記比較ステップにおける比較結果に基づき、前記第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定するステップと、

を有する方法。

【請求項２】　前記比較ステップは、第１のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第１の系列と、第２のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第２の系列との間の差を算出するステップを有し、

前記判定ステップは、前記算出された差に基づき、判定することを特徴とする、請求項１に記載の方法。

【請求項３】　さらに、第１のコネクションにおけるパケットのデータ・サイズと検知時刻との情報を含む、第１のパケット・データを受信するステップと、

前記受信した第１のパケット・データに含まれる検知時刻に基づき、比較するパケット・データを検索するステップと、

前記検索ステップにおける検索結果に基づき、前記第２のコネクションに含まれるパケットを選択するステップと、

を有する、請求項１又は２に記載の方法。

【請求項４】　前記検知時刻はパケット・データに含まれるタイム・スタンプによって特定され、前記データ・サイズはシーケンス番号によって特定される

、請求項１、２又は３に記載の方法。

【請求項５】　前記比較ステップは、前記第２の系列の部分列であって、最初の項をずらすことによって形成される複数の部分列と、前記第１の系列とを順次比較するステップを含む、請求項１、２、３又は４に記載の方法。

【請求項６】　第１のコネクションにおけるパケットのデータ・サイズと、パケットの検知時刻とを含む第１のパケット・データを記録するステップと、

第２のコネクションにおけるパケットのデータ・サイズと、パケットの検知時刻とを含む第２のパケット・データを記録するステップと、

前記記録した第１のパケット・データを送信するステップと、

前記送信された第１のパケット・データを受信するステップと、

前記受信した第１のパケット・データと前記第２のパケット・データに基づき、前記第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、前記第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較するステップと、

前記比較ステップにおける比較結果に基づき、前記第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定するステップと、

前記判定ステップにおける判定結果を、送信するステップと、

を有する、方法。

【請求項７】　コンピュータ読み取り可能な記録媒体であって、請求項１、２、３、４、５又は６に記載した方法の処理を、コンピュータに行わせるプログラムを記録した、記録媒体。

【請求項８】　複数のコネクションから構成されるアクセス・チェーンを介して、ネットワーク上をパケットが通信されるシステムにおけるアクセス・チェーンを追跡するためのシステムであって、

第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較する比較部と、

前記比較部における比較結果に基づき、前記第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定する判定部と、

を有するシステム。

【請求項９】　前記比較部は、第１のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第１の系列と、第２のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第２の系列との間の差を算出し、

前記判定部は、前記算出された差に基づき、判定することを特徴とする、請求項８に記載のシステム。

【請求項１０】　さらに、第１のコネクションにおけるパケットのデータ・サイズと検知時刻との情報を含む、第１のパケット・データを受信する受信部と、

前記判定部における判定結果を送信する送信部と、

を有する、請求項８又は９に記載のシステム。

【請求項１１】　前記受信した第１のパケット・データに含まれる検知時刻に基づき、比較するパケット・データを検索する検索部と、

前記検索部における検索結果に基づき、前記第２のコネクションに含まれるパケットを選択する選択部と、

を有する、請求項１０に記載のシステム。

【請求項１２】　前記検知時刻はパケット・データに含まれるタイム・スタンプによって特定され、前記データ・サイズはシーケンス番号によって特定される、請求項８、９、１０又は１１に記載の方法。

【請求項１３】　前記比較部は、前記第２の系列の部分列であって、最初の項をずらすことによって形成される複数の部分列と、前記第１の系列とを順次比較する、請求項８、９、１０、１１又は１２に記載の方法。

【請求項１４】　複数のコネクションから構成されるアクセス・チェーンを介して、ネットワーク上をパケットが通信されるシステムにおけるアクセス・チェーンを追跡するためのシステムであって、

パケット・サイズと検知時刻の情報を含むパケット・データを記録する記録部と、

前記記録したパケット・データを、判定のために他のサイトに送信する送信部

と、

前記サイトからの判定結果を受信する、受信部と、

を有する、システム。

【請求項１５】

パケットのデータ・サイズと検知時刻とを含む第１のパケット・データを収集し、前記第１のパケット・データを送信する、第１の収集装置と、

パケットのデータ・サイズと検知時刻とを含む第２のパケット・データを収集する、第２の収集装置と、

前記受信した第１のパケット・データと前記第２のパケット・データに基づき、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較し、前記比較結果に基づき、前記第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定する、計算システムと、

を有するネットワーク・システム。

【請求項１６】

前記計算システムは、第１のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第１の系列と、第２のコネクションにおけるパケットのデータ・サイズと検知時刻とに基づいて特定される第２の系列と、の間の差を算出し、

前記前記算出された差に基づき判定する処理を行う、請求項１５に記載のネットワーク・システム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

この発明は、ネットワークにおけるアクセス・チェーンの追跡技術に関するものであり、特に、パケットのデータ・サイズと検知時刻に基づき、アクセス・チェーンを追跡する技術に関するものである。

【０００２】

【従来の技術】

インターネットの利用は時間・空間を越えた情報アクセス手段の提供およびその匿名性などから急速に広がった。ネットワーク犯罪もインターネットを利用するので、アタッカーにも同様の利便性を与えてしまった。このためアタッカー及びアタック・ターゲットは広域化し、アタッカーを特定することや攻撃内容を把握することが困難になってしまった。またアタック・ツールがホームページを通して提供されることで攻撃手法が一般に広まってしまい、不正アクセスが大量発生する恐れがでてきた。

【０００３】

一方、不正アクセスから守るべきシステムは日々変化している。例えばLinuxではパッケージのリリースアップが頻繁に行われ、弱点を塞ぐためのアップデート・モジュールが月に１、２個提供されている。そのうえ、上記のように、不正アクセスの手法は高度化し相手を特定することが困難になってきている。この状況では何らかの防御システムで不正アクセスを完全に防ぐことは難しいと言わざるを得ない。上記の理由から、不正アクセスの大量発生を抑止するシステム作りが急務である。

【０００４】

ネットワークを利用した犯罪にはメール爆弾、使用不能攻撃、侵入、誹謗中傷などがあるが、これらの犯罪ではアタッカーが身元を隠すために第三者のシステムから攻撃する事が多い。この第三者のシステムのことを踏み台コンピューターと呼ぶ。

【０００５】

この不正侵入の方法について、図１を用いて説明する。図１は、複数のホストコンピュータ１２～１５と、これらホストコンピュータが接続されているネットワーク１７を介して、アタッカー・コンピュータ１１が、ターゲット・ホスト１６に不正侵入する様子を、概略的に示したものである。ネットワーク１７上のパケットは、ルータ１８によって経路が選択される。アタッカー１１は身元を隠すために一つ以上の踏み台コンピューター１２～１５を経由して攻撃を仕掛ける。

【０００６】

アタッカー１１は自由に利用できる踏み台コンピューターを得るために、シス

テム設定の不備やOSのバグ等を利用して侵入を試みる。アタッカー１１がターゲットとするコンピュータ１６に攻撃する際は、いくつかの踏み台コンピューター１２～１５を経由し、最終的にはターゲットのコンピュータ１６に対して攻撃パケットを送出する。このアタッカー１１からターゲットコンピュータに至るコンピュータ１６の連鎖を不正アクセス・チェーンと呼ぶ。不正アクセス・チェーンはtelnetやrloginを使用することが一般的ではあるが、ポート番号を変更したサービスを使用する場合もある。

【０００７】

現在のインターネット環境では、不正アクセス・チェーンの追跡をすることは簡単な問題ではない。まず、パケットの送信元アドレスからは一つ前の踏み台ホストのIPアドレスしかわからない。よってその踏み台ホストより前のホストの送信元アドレスを調査したくとも、その踏み台ホストは第三者により管理されているため追跡の解析は許可されないことが一般的である。したがって不正アクセスに関わるパケットを洗い出して、その送信元アドレスを順次たどって逆探知をするのは困難である。

【０００８】

そこでシステム的に不正アクセス・チェーンを自動追跡する手法が研究されている。不正アクセス追跡の手法は、追跡を行うためのコンポーネントをどこに設置するかによって、大きく分けて「ホストベース」と「ネットワークベース」の２つに分類できる。ホストベースの手法は、追跡をするためのコンポーネントを各ホストに設置するのに対して、ネットワークベースの手法は追跡をするためのコンポーネントをネットワークのインフラ（例えばルータやスイッチなど）に設置する。ホストベースベースの追跡の手法としては、以下のようなものが考えられている。

【０００９】

DIDS (Distributed Intrusion Detection System)

UC Davisで最初作られ、その後Trident Data Systems社に引き継がれたこのシステムは、管理下のネットワーク内で起こるすべてのTCPコネクションとログインを監視する。それによって、ユーザのログインによる移動などの行動や現状を

常に把握する。管理下の各ホストにはそれぞれホストモニターが動いており、自分のところで起こるネットワークアクセスの監査情報を収集して、中央のDIDSディレクターへ送信し、そこで管理下のネットワーク状況が集中管理される。この技術内容については、"S. Snapp et al. DIDS (Distributed Intrusion Detection System) - Motivation、 Architecture、 and An Early Prototype In Proceedings of the 14th National Computer Security Conference、 1991"に詳細に説明されている。

【００１０】

CIS (Caller Identification System)

ログインの際に通信元を確認する手法である。N-1個のホストを辿ってN番目のホストにログインする際に、N-1番目のホストにそれ以前のホストのリストを問い合わせる。また１からN-2番目のそれぞれのホストにそれ以前のホストのリストを問い合わせる。問い合わせ結果に食い違いがないことを確認してログインを許可する。この手法は管理されたホスト間のアクセスを許可するためのもので、管理下にあるホストにはCISを導入することを前提とする。この技術内容については、"H. T. Jung et al. Caller Identification System in the Internet Environment、 Proceedings of the 4th Usenix Security Symposium、 1993"に詳細に説明されている。

【００１１】

Tsutsuiの手法

この手法は米国特許5、220、655号公報に開示されており、各ホストのファイルシステムにそのホストにアクセスしたユーザとプロセスの情報を保管しておき、追跡要求があったときにそれらの情報を順次取得して辿っていくことでユーザのアクセス・チェーンを得るというものである。もし追跡要求が異なる管理ドメインに入る場合はそのドメインの管理ホストが必要な情報を収集してから要求元に返す。各ホストではトラッキングサービスプロセスが動いていなければならない。

【００１２】

平田らの手法

特開平10-164064号公報に開示されたものであり、この手法は各ホストが自分に関わるコネクションのプロセスとポート番号をアクセスログ記録部に保管しておき、基本制御プログラムがホスト間のアクセス情報を交換してアクセス・チェーンを辿るというものである。基本制御プログラムが経路追跡に必要な処理を行うので、アプリケーションプロセスは追跡に関する制御を全く意識する必要がない。

【００１３】

以上のようなホストベースの手法で問題なのは、アクセス・チェーンの途中のあるホストがそのシステムを導入していなかった場合、追跡はその地点まででそれ以上は完全に不可能となってしまう点である。インターネットではある特定のホストベースのシステムを全ての管理ドメインのホストが採用するということは考えにくい。また、管理下のホストであったとしてもそのホストが侵入されて追跡に関わるプログラムが書き換えられていることも考えられる。したがってインターネット環境でホストベースのシステムを高い信頼性で実行させることは現実的ではないといえる。

【００１４】

この他に、アメリカ海軍により報告されたシステムで、不正アクセスにより踏み台とされたホストを逆向きに辿ることで行為者を発見する手法である、Caller IDがある。この追跡方法は不正アクセス行為者と同様の方法で踏み台とされたホストを逆順に侵入していくというものである。探索側は踏み台とされたホストに侵入しなくてはならないが、すでに侵入された事実から同様に侵入できると主張している。実際の追跡では侵入は困難であったり、不正アクセス行為者によるセキュリティホールの修正などで侵入は不可能となる場合もある。また第三者のコンピュータに侵入することはあらたな犯罪とも考えられる

【００１５】

次に、ネットワークベースの手法を説明する。

Staniford-Chenらの研究では通信データの内容のみに着目し、データの特徴量(通信文字種の分布)は個々のセッション（侵入行為）ごとにユニークであり、この特徴量はアクセス・チェーンのどのコネクションにおいてもほとんど同じであ

るという仮定で不正アクセス・チェーンを見つける。ネットワーク上のできるだけ多数の地点（ルータなど）で、各セッションに対して一定時間間隔毎に区切ってこの特徴量をそれぞれ計算して保管しておく。もし侵入が発覚した場合、侵入に使われたセッションの特徴量と近い特徴量を持つものをネットワーク上の多数の地点で探すことでアクセス・チェーン上にあったホストを発見する。この技術内容については、”Stuart Staniford-Chen and L. Todd Heberlein、 Holding Intruders Accountable on the Internet、 Proceedings of the 1995 IEEE Symposium on Security and Privacy、 Oakland、 CA.、1995”に記載されている。

【００１６】

この手法の利点は通信データの特徴量を計算するだけのログを保管すればよいので少ない容量と軽い処理でログ保管できること、また逆探知をする上で順番に辿る必要はないということである。しかしながら、通信データの内容に依存するため、暗号処理や言語コードの置換などによりデータ内容が変更された場合に、全く対処できない。

【００１７】

尚、他の技術としては、特開平9-2114493に開示されているように、ネットワーク内の計算機の監査システムであって、計算機の監査結果に以上があった場合に、ログ収集装置においてその計算機のトラヒック・ログを収集する技術がある。

【００１８】

【発明が解決しようとする課題】

追跡手法のうちホストベースのものはインターネットのような多用な管理権限をもつネットワークでは利用しづらい。一方、ネットワークベースのThumbprintingは有効性が高い方法ではあるが、暗号化通信が普及してきている現在、データ内容を元に照合するこの手法をいくら発展させても、アクセス・チェーンの発見手法としては将来性が低いといえる。

【００１９】

そこで、本発明の一つの目的は、ネットワークで部分的にでも取り入れられていれば、その範囲内で部分的にアクセス・チェーン上のホストを見つけることが

できる、アクセス・チェーンの追跡方法及びそのシステムを提供することである。

【００２０】

本発明の他の目的は、パケットのデータ内容に依存せずに、パケットのデータ内容が、途中の経路で暗号化されたり言語コードが変換されたりする場合にも対応することができる、アクセス・チェーンの追跡方法及びそのシステムを提供することである。

【００２１】

本発明の他の目的は、見つかったアクセス・チェーン上の複数のホストのうち、どれが最もアタッカーに近いものであるかが容易に分かる、アクセス・チェーンの追跡方法及びそのシステムを提供することである。

【００２２】

本発明の他の目的は、アクセス・チェーンの追跡に必要なパケットのデータを記録するときに、必要とする記憶容量が少なくてすむ、アクセス・チェーンの追跡方法及びそのシステムを提供することである。

【００２３】

本発明の他の目的は、パケットのデータの内容を記録することなく、通信におけるプライバシーの保護をも図ることができる、アクセス・チェーンの追跡方法及びそのシステムを提供することである。

【課題を解決するための手段】

本発明の一つの態様は、パケットのデータ内容に依存せずに、パケットのデータ・サイズと処理時刻に基づき、アクセス・チェーンの追跡を行う。

【００２４】

本発明の他の態様は、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較し、この比較結果に基づき、前記第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定する、アクセス・チェーンの追跡方法である。

【００２５】

本発明の他の態様は、第１のコネクションにおけるパケットのデータ・サイズと、パケットの検知時刻とを含む第１のパケット・データを記録するステップと、第２のコネクションにおけるパケットのデータ・サイズと、パケットの検知時刻とを含む第２のパケット・データを記録するステップと、記録した第１のパケット・データを送信するステップと、送信された第１のパケット・データを受信するステップと、受信した第１のパケット・データと第２のパケット・データに基づき、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、前記第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較するステップと、比較ステップにおける比較結果に基づき、第１のコネクションと第２のコネクションとが、同一のチェーンに含まれるかを判定するステップと、判定ステップにおける判定結果を、送信するステップとを有する、アクセス・チェーンの追跡方法である。

【００２６】

本発明の他の態様は、複数のコネクションから構成されるアクセス・チェーンを介して、ネットワーク上をパケットが通信されるシステムにおけるアクセス・チェーンを追跡するためのシステムであって、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、第２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較する比較部と、比較部における比較結果に基づき、第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定する判定部とを有するシステムである。

【００２７】

本発明の他の態様は、複数のコネクションから構成されるアクセス・チェーンを介して、ネットワーク上をパケットが通信されるシステムにおけるアクセス・チェーンを追跡するためのシステムであって、パケット・サイズと検知時刻の情報を含むパケット・データを記録する記録部と、記録したパケット・データを、判定のために他のサイトに送信する送信部と、サイトからの判定結果を受信する、受信部とを有する、システムである。

【００２８】

本発明の他の態様は、パケットのデータ・サイズと検知時刻とを含む第１のパ

ケット・データを収集し、前記第１のパケット・データを送信する、第１の収集装置と、パケットのデータ・サイズと検知時刻とを含む第２のパケット・データを収集する、第２の収集装置と、受信した第１のパケット・データと第２のパケット・データに基づき、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、２のコネクションにおける時刻によるパケットのデータ・サイズの変化と、を比較し、比較結果に基づき、第１のコネクションと前記第２のコネクションとが、同一のチェーンに含まれるかを判定する、計算システムと、を有するネットワーク・システムである。

【００２９】

本発明の上記及びその他の態様は、以下の発明の実施の形態及び図面の記載によって開示される。

【００３０】

【発明の実施の形態】

本実施形態は、TCP/IPネットワーク上の複数のホストに連鎖的にtelnetやrlogなどのログインを繰り返して、リモート・ホストを操作している侵入者がいるとき、この連鎖としてのアクセス・チェーンの途中において、ネットワーク上の複数の位置でそれぞれ行き交うパケットのログを記録することで、その侵入ルートを、パケットのヘッダ情報と検知時刻の情報に基づいて判別する方法について説明する。

【００３１】

ネットワーク上の不正アクセス・チェーンによる侵入.

図２は、本実施形態の、ネットワーク・システム全体を概略的に示す構成図である。図において、２１はバックボーン（又はバックボーン・ネットワーク）である。バックボーン２１とは、ネットワーク の基幹回線、または基幹ネットワークのことで、ＦＤＤＩやＡＴＭなど高速通信可能なネットワークをルータやスイッチなどで繋いで構成される。２２は、バックボーンに接続されている各ネットワークであり、例えば、ＮＳＰ（ネットワーク・サービス・プロバイダ）がそれぞれ独自にもつネットワークである。２３はネットワーク上に設置されている各ルータ、２４は各ネットワークに設置されているコンピュータ計算機である。

２５は、ネットワークを通信されるパケットのログ情報を収集する、ログボックスであり、ネットワークに接続されている。

【００３２】

図３は、図２に示したネットワークにおいて、アタッカーAがネットワーク上の何台かのホストへ次々と連鎖的にtelnetやrloginを繰り返し、これらの踏み台ホストを経由して、最終的にはターゲットTを攻撃している状況を概念的に示した図である。このホスト間のコネクションの連鎖を不正アクセス・チェーンと呼ぶ。図３において、３１はアタッカー、３２～３５は踏み台ホスト、３６はターゲット・コンピュータ、３７～３９はログ・ボックスである。４０はルータであり、本来は複数のルータを経由して侵入が行われるが、便宜的に、一つのバックボーン・ルータのみを図示している。バック・ボーンルータとは、バックボーン・ネットワークに接続されているルータである。踏み台ホストＥと、踏み台ホストＤとの間を行き交うパケットは、途中でルータ４０を経由して流れている。

【００３３】

アタッカー３１は、ホストへ次々と連鎖的にtelnetやrloginを繰り返して、不正侵入を行う。このとき、アタッカー３１は、キー入力によって対話的に踏み台となるリモート・ホストを操作している。telnetやrloginの通常の使用方法では，キー入力に対応するパケットは，データが複数のパケットに分割して入ったり、複数のパケットの中のデータが一つのパケットの中に統合されたりしながらも、アタッカーＡから踏み台ホストＢ３５まで連鎖を経由して確実に流れる。コマンドの実行結果など、リモートコンピュータで画面に表示される内容に対応するパケットは、アクセス・チェーンを逆向きに同様に流れる。

【００３４】

telnetやrloginの通常の使用方法では、パケットの中のデータ部分は、分割や統合などがあっても合計としては連鎖の途中で一般的に一定である。すなわち、コマンドをキー入力したら、それが変化せずにそのまま踏み台の連鎖を伝わり、コマンドの実行結果は同じく変化せずにそのまま踏み台の連鎖を逆向きに伝わる。向きとは、全てのパケットは連鎖の途中のどの位置においてもアタッカーA３１から踏み台B３５への向きと、それと逆向きの２種類に分けられるということ

である。尚、データ部分のサイズの合計が一定とは、連鎖の途中でパケット内のデータ部分に余分なデータが追加されたり、余計なパケットが新たに作りだされたりしないということを意味する。

【００３５】

パケット・データの記録装置の設置.

次に、パケット・データの記録装置としてのログ・ボックスの設置方法について説明する。パケット・データの記録は、インターネット上の可能な限り多くのしかも様々な場所で行うことが望ましい。侵入者のアクセス・チェーンの一部が、パケットデータを記録していたいづれかの場所を通過していた場合に、その部分を検出できるため、データを記録している場所が多ければ多いほど、侵入者が使ったアクセス・チェーンをより完全な形で発見できる可能性が高まるからである。

【００３６】

パケットのログ・データを記録するインターネット上の場所には、バックボーン２１が適している。インターネットの様々なＮＳＰ（ネットワーク・サービス・プロバイダ）が相互に接続されているポイントでパケット・データを記録すれば、それらＮＳＰのネットワーク間を行き交う全てのパケットを捕らえることができる。図２においては、ログ・ボックス２５は、バックボーン２１とＮＳＰ２２のコネクション部分に設けられていることが、示されている。また各ＮＳＰのネットワーク内に閉じた範囲を行き交うパケットは、そのＮＳＰのバックボーン内でパケットを記録することができる。従って、これらバックボーンを管理する組織が、バックボーン内にパケット・データを採取するＰＣであるログ・ボックス２５を設置することによって、侵入者追跡を行うことが可能となる。

【００３７】

パケット・データの記録のための記録手段として使用されるログ・ボックスは、本実施形態においては、大容量のＨＤＤと、ネットワーク・カードを備えたＰＣである。ネットワーク・カードは、送受信装置としてのインターフェースとして機能する。もちろんＰＣを用いずに、専用の装置を新たに構成してもよい。このＰＣ２５は、ルータの１つのコネクタに接続される。図４はこのログ・ボック

ス２５のハード・ウェア構成を示したものである。図４において、４１はパケット・データの入出力を制御する入出力制御部としてのネットワーク・カード、４４はフィルタ・プログラムやパケットのログ・データを格納する記録装置としてのＨＤＤ、４２は各種処理及び制御を行うＣＰＵ、４３はＣＰＵの処理において利用される一時記憶装置としてのＲＡＭである。

【００３８】

次に図５は、ログ・ボックス２５のソフト・ウェア構成を示している。ログ・ボックス２５は、データ受信制御情報の格納、外部との送受信を制御する送受信制御部５１、パケットから必要な情報を収集するための情報を含み、フィルタの部の動作を制御するフィルタ・プログラム５２、フィルタ・プログラムに従い受信したパケットから必要なパケット情報を選択するフィルタ部５３、フィルタ部５３が選択収集したデータを記録する記録部５４、装置全体の管理・制御を行うシステム制御部５５を有している。

【００３９】

動作について説明する。ネットワークを通信されるパケットが、ルータからログ・ボックスに送信される。ログ・ボックス２５の送受信制御部５１がこのパケットを受信し、フィルタ部５３に送る。フィルタ部５３は送られたパケットから、フィルタ・プログラム５２に従い、必要なパケット・データを取得する。このパケット・データは、ログ・データとして記録部５４に記録される。全体の動作はシステム制御部５５によって制御される。尚、具体的なパケット・データの取得方法については、後に記載する。

【００４０】

ログ・ボックス２５のネットワークへの接続方法について説明する。ログ・ボックス２５のネットワーク・インターフェース・カードはＰＣでよく用いられているイーサネット・カードであり、ルータ２３はイーサネットをネットワーク・インターフェースに持つ。ルータ２３のあるイーサネットのコネクタを一つ決め、ルータ２３を通過する全てのパケットのコピーを本来のルーティング先に加えてそのイーサネットのアドレスへも送信するように設定を変更する。ルータ２３のそのイーサネットのコネクタとログボックス２５のイーサネット・カードのコ

ネクタをケーブルで結ぶ。こうすることでログボックス２５はルータ２３を通過する全てのパケットのコピーを受け取ることができる。

【００４１】

パケット・データの記録

必要なパケット・データの記録方法について説明する。本実施形態においては、ＴＣＰパケットのみを記録する。もちろん、他のデータ形式を選択するようにしてもよい。まず、このＴＣＰパケットの構造について説明する。ＴＣＰ／ＩＰの通信では、データは基本的にはＩＰパケットとして構成される。ＩＰパケットは　[ＩＰヘッダ] [ＩＰデータ]という並びのデータで構成される。次にＴＣＰ通信の場合（ＩＰヘッダのProtocolの部分がＴＣＰを表している場合）、[ＩＰデータ]は　[ＴＣＰヘッダ] [ＴＣＰデータ]という並びのデータで構成される。

【００４２】

ＩＰヘッダ の構造を図６を用いて説明する。この図の横軸はビットであり、３２ｂｉｔ（４ｂｙｔｅ） 毎に改行して下の行の左端に続いている。同一の行においては、左側のビットがより上位のビットを表している。Options が無い通常のＩＰヘッダは Version から　Destionation Address までの２０byteとなる。ここで、ソース・アドレス（Source address）とデスティネーション・アドレス（Destination address）は、それぞれ、送信側の装置のＩＰアドレスと送信先（受信）の装置のＩＰアドレスである。次に、図７はＴＣＰパケットの構造を示している。構造の示し方は、図６と同様である。Options とdata を除いた通常のＴＣＰヘッダはSource PortからUrgent Pointerまでの２０byteとなる。ここで、ソース・ポート（Source port）、デスティネーション・ポート（Destination port）、シークエンス・ナンバ（Sequence Number）とは、それぞれ、送信側の装置のポート番号、送信先（受信）の装置のポート番号、そして、一つのコネクションにおける各パケットに与えられる番号である。尚、これらデータ構造は周知のものであり、ここでは詳細な説明は行わない。

【００４３】

所望のパケット・データを記録するには、まず、ネットワーク・カード４１に届いた全てのパケットから所望のパケットだけを選別する。そして、指定した最

大長までパケットを切り取り、届いた時間順にそれをＨＤＤに記録していく。この処理を、ソフト・ウェア構造を使用して、具体的に説明する。ルータから送られたパケット・データを送受信制御部５１が受信する。フィルタ部５３は、この受信パケットの内、フィルタ・プログラム５２に従い、必要な種類のパケットのみを選択する。本形態では、ＴＣＰパケットのみを選択する。フィルタ・プログラム５２は、予め記録すべきパケットのデータ・サイズが記録されている。フィルタ部５３はこのデータ・サイズに従い、選択したパケットの先頭からこのデータ・サイズ分を取得し、所望の処理をして記録部５４に記録する。

【００４４】

記録部５４に書きこまれる各パケット・データの概略構造を、図８に示す。Time Stampはパケットが送受信制御部５１に受信された時刻である。caplenは、受信パケットから切り取られたデータ・サイズ（byte)である。lenは実際にネットワーク・カードが受信したパケットのデータ・サイズ（長さ（byte））である。data部には、実際に取得されたパケットのデータ（バイト列）がcaplen（byte)のデータ・サイズで書き込まれる。記録部５４に書きこまれるファイルは、先頭にフィルタ・プログラム５２のバージョン情報などのヘッダー部分が入り、その後のデータは、上記フォーマットのデータが時間順に連続して記録される。

【００４５】

各パケットから切り取るデータの長さは、ＩＰヘッダーとＴＣＰヘッダーが含まれるのに十分に長さにすればよい。イーサネットではイーサネットのヘッダーが１４バイト、その後にＩＰヘッダが通常２０バイト、その後ＴＣＰヘッダが通常２０バイトであるのでＩＰヘッダやＴＣＰヘッダにオプションが入っていて多少長くなったとしてもそれで十分なサイズとしては、６８バイトぐらいが適当である。ある程度の数のパケットを一つのファイルに書き出したらいったん終了して、また別のファイルに続きを書き出す処理を行う。これを繰り返して、ＨＤＤ４４が一杯に近づいたら、最も古いファイルから上書きする。こうすることで常に過去何日か分のパケット・データを蓄積している状態にする。尚、フィルタ・プログラム５２としては、パケット・キャプチャ・ソフトを利用することも可能である。

【００４６】

アタッカー侵入の発見.

あるシステム管理者が、自分の管理下のコンピュータがアタッカーによる侵入を受けたことに気づいた場合の処理を説明する。これは自分のシステムが、直接攻撃対象になっていた場合もありえるし、もしくは、他のシステムを攻撃するための踏み台として侵入を受けた場合もありえる。いづれにせよ、侵入に使われたコネクションのパケット・データが記録されていなければならない。たとえそのときは記録していなかったとしても、アタッカーは大抵侵入したコンピュータに何らかのバック・ドアを仕掛けておき、後日再び同じコンピュータへ侵入することがしばしばある。そのため、そのコンピュータと同一ＬＡＮ内の他のコンピュータにおいて、その後のパケット・データを常に記録するようにし、再びアタッカーが侵入してくるのを待てば良い。

【００４７】

尚、イーサネットを使ったＬＡＮの場合には、同一ＬＡＮ内（ルータを経由しないでハブ（シェアード・ハブ）で繋がっている範囲）の全ネットワーク・トラフィックは、そのネットワーク内のどれか一台のコンピュータで見ることができる。

【００４８】

前回の侵入時のパケット・データを記録していなかった場合は、その後、再びアタッカーが侵入してきたとき、その間のパケット・データを記録する。ただし、ここで記録されたパケット・データには、アタッカーとは関係ないコネクションも含まれているため、アタッカーが侵入に使ったのが、どのコネクションだったのかを判断する必要がある。侵入を受けたコンピュータのログは、多くの場合アタッカーによって書きかえられているため当てにならない。従って、ログ・ボックスに記録したアタッカーの侵入時間、ソースＩＰアドレス、ソース・ポート番号などから、どのコネクションだったのかを判断する。多くの場合、当人が全く身に覚えのない時間帯と場所から、その人のユーザＩＤでのログインがあったり、長い間使われていなかったユーザＩＤでのログインがあったときなど、それは不正侵入である可能性が高い。

【００４９】

各協力サイトへのパケット・データの配布.

不正侵入が発見されると、アタッカーが使ったコネクションのパケット・データが、ネットワーク上のログボックスを設置した各協力サイトに配信される。サイトとの間では、データ転送に先立って、お互いのコンピュータ同士で、予め認証が済んでいるものとする。

【００５０】

ログ・ボックスは、記録部５４に記録されているパケット・データから、所望のパケット・データを選択する。基本的には、パケットのＩＰヘッダとＴＣＰヘッダから得られるコネクションを同定する４つ指標（ソース・ＩＰアドレス、ソース・ポート番号、デスティネーションＩＰアドレス、デスティネーション・ポート番号）から、コネクションを特定することができるので、各パケット・データが、アタッカーが使用したコネクションに含まれるかどうかを判別できる。従って、それらが一致するパケット・データをファイルに書き出す。データ・ファイルに書き出されるパケット・データは、記録部が記録したパケット・データと同じ形式である。データ・ファイルは、探索要求パケットと共に、各サイトに配信される。配信された要求パケットとデータ・ファイルは、そのサイトの各ログ・ボックスが受信する。要求パケットを受信した各ログ・ボックスは、その比較判定プログラムを起動する。

【００５１】

ここで、図９を用いて比較判定装置のソフト・ウェハ構成を説明する。図において、９１はネットワークとの間でのデータの送受信を制御する送受信制御部、９２はコネクションの比較、類似度の判定を制御する比較判定プログラム、９３はパケット・データ等を記録する記録部、９４は全体の制御を行う制御部、９５は比較判定プログラム９２に従い比較判定処理を行う比較判定部である。尚、本実施形態においては、ネットワーク・カードを備える計算機であるログ・ボックスが、パケット・データの収集と、コネクションの比較の双方を行うが、もちろん、他の計算機がこのコネクションの比較判定処理を行ってもよい。又、本実施形態においては、パケット・データを送信しているが、このパケット・データに

所望の処理を行い、例えば、比較判定システムで使用される系列データに変換してから送信してもよい。

【００５２】

各サイトの処理.

各協力サイトに、侵入を受けたサイトからアタッカーの使用したコネクションのパケット・データ・ファイルが送られてきた場合、各ログ・ボックスにおいて行われる処理の要約は、以下のようなものである。まずこのファイルの先頭と最後のパケットの時刻を調べることで、蓄積されたデータのどの時刻からどの時刻までの部分を用いて検索すればよいかを決める。該当する時間の範囲にある蓄積されたファイルを使ってアタッカーの使用したコネクションの時刻-データサイズのパターンと類似しているものを見つける。見つかった類似コネクションにはそれぞれ類似度をあらわす点数がつけられ、上位のいくつかのコネクションが侵入を受けたサイトへ送り返される。

【００５３】

類似コネクションの特定法.

２つのコネクションの類似を判定する方法について説明する。一つのコネクションには互いに逆向きの二方向のデータの流れがあるが、それぞれについて「時刻－シーケンス番号」系列を考える。これは、そのコネクションの一方向のデータの流れについて、パケットが通過した時刻とそのパケットのシーケンス番号を対にした系列である。「シーケンス番号」は、ＴＣＰのコネクションが確立する毎に最初ランダムな値が決まり、その後は、データグラム・パケットが到着する度にデータ量（バイト）ずつ増えていく番号である。ここで「データグラム・パケット」とは、データ部分のサイズが０より大きい（ヘッダのみではない）ＴＣＰパケットを言う。パケットのヘッダには、そのパケットのデータの先頭のバイトのシーケンス番号が入っているので、これにデータ部分のサイズを足せば、データの最後のバイトのシーケンス番号が分かる。以下ではパケットのシーケンス番号はデータ部分の最後のバイトのシーケンス番号を示すとする。

【００５４】

この系列の具体例を以下に示す。例えば、9.116.158.27:23 --> 9.116.77.22

5:40509 というコネクションの一方向のデータの流れについて、初期シーケンス番号が940000だったとする。「時刻－ シーケンス番号」系列は以下のような系列になる。

【００５５】

時刻	シーケンス番号
14:49:04.026199	940003
14:49:04.140934	940054
14:49:04.305649	940087
14:49:04.372342	940094
14:49:04.462903	940097
14:49:05.731234	940098
14:49:05.860761	940099
14:49:06.372575	940100
14:49:06.439558	940101
14:49:06.698750	940102
14:49:06.773162	940103
・・・・・・・	・・・・

【００５６】

時間の横軸、シーケンス番号の縦軸において、最初のパケットの時刻と初期シーケンス番号を原点として、系列をグラフにあらわしたものが、図１０に示されている。あるコネクションの一つの方向の「時刻－シーケンス番号」系列が与えられたとき、それと類似する系列を多数の他の系列の中から探すことを考える。

類似度の計算には様々な方法が考えられる。一般的に２つの実数の系列 $\{x_1, x_2, \cdots, x_n\}$、$\{y_1, y_2, \cdots, y_n\}$ の類似度を考える問題では、これらをn次元空間の点とみてその距離を

$$(\Sigma |x_i - y_i|^p)^{(1/p)}$$

で定義して、この値が０に近いほど類似度が高いと考えられる。p=2 のときはよく知られているユークリッド空間での距離である。

【００５７】

類似度を与える方法の一例として、２つの系列をグラフで表現したとき、それらを近づけると、どのくらいその形状が一致するかという指標が考えられる。２つのグラフの間に挟まれる領域の面積を縦軸方向（シーケンス番号）の範囲の長さで割った値、つまり２つのグラフを近づけると平均してどのくらいの横軸方向（時間）の隙間が生じるかを類似度の指標にとることにする。この値が０に近いほど２つの系列の類似度は高い。この計算は２つのパケット系列の時刻の差を各シーケンス番号について総和して、これをシーケンス番号の範囲の長さで割ればよい。ある踏み台コンピュータでのパケット系列をＡ、これと比較するパケット系列をＢとする。見つけたいパケット系列Ｂはアクセス・チェーンにおいて踏み台より上流（アタッカーにより近い側）の系列であり、系列Ｂのデータ量は系列Ａのデータ量以上であるので、データ系列Ｂのグラフの縦軸方向の範囲は系列Ａの縦軸方向の範囲を含んでいなければならない。また縦軸方向に関して系列Ａの開始位置は系列Ｂの途中のどこかの位置に対応している。またグラフをお互いに近づける場合、ＢをＡと交わらない位置まで横軸方向へ平行移動させて考える。したがって系列Ｂのグラフを、系列Ａのグラフと交わらず縦軸方向の範囲で含むようにして、上下左右に、ＡとＢが挟む領域の面積が最小になるような位置に移動して、そのときの２つのグラフに挟まれる面積をＡのシーケンス番号の範囲の長さで割った値を２つの系列の類似度とする。

【００５８】

このようにして、与えられたコネクションの一つの方向の系列と全てのコネクションのそれぞれ２つの方向の系列との類似度を計算して、類似度の高いものを探し出す。

【００５９】

類似度を与える他の方法としては、全ての系列を離散フーリエ変換して、時間についての系列から周波数についての系列に変換することが考えられる。この場

合は系列としてデータグラム・パケットの通過時刻とデータ量（累積ではない）とを利用する。得られたフーリエ係数の最初の何項かのみで、時間軸ベースに戻したときのその系列の特徴をほぼ表しているため、比較する系列の次元を減らすことができる。このようなフーリエ変換を用いる手法は様々なバリエーションが考えられる。

【００６０】

系列間面積の算出.

２つの系列の差を算出する具体的な方法について説明する。アタッカーの使用したコネクションのパケットデータＡが書かれているファイルを基に、特定のサイトのログ・ボックスに記録されたパケット・データＤの中の、一つのコネクションとの類似度を計算する手法を説明する。これは、上記の面積の算出を具体的に説明している。以下の計算の理解のために、図１０を参照して欲しい。

【００６１】

パケットqのタイムスタンプ（時刻）をT(q)、パケットqのデータの最後のバイトのシーケンス番号をS(q)と表す。またパケット系列 $Q = \{ q_1, q_2, \ldots, q_n \}$ のグラフは２次元平面で、横座標が $T(q_i)$、 縦座標が$S(q_i)$ の階段状のグラフとする。つまり Q の初期シーケンス番号を Sq とすると、

点 $(T(q_1), Sq)$, $(T(q_1), S(q_1))$, $(T(q_2), S(q_1))$,$(T(q_2), S(q_2))$, ..., $(T(q_i), S(q_{i-1}))$, $(T(q_i), S(q_i))$, ..., $(T(q_n), S(q_{n-1}))$, $(T(q_n), S(q_n))$

を結んだグラフである。

【００６２】

１） パケットデータＡが書きこまれているファイルからコネクションの二方向のうち一方向を指定して、その方向のパケット系列P

$P (=P(n)) = \{ p_1, p_2, \ldots, p_n \}$

を作成する。系列P を参照としてこれとの類似度をＤの中の各パケット系列について求める。

【００６３】

２） パケット系列 $P(n) = \{ p_1, p_2, \ldots, p_n \}$ のグラフの始点を原点にそ

ろえたグラフをP'(n) とする。

$P'(n) = \{ p'_1, p'_2, \ldots, p'_n \}$,
$T(p'_k) = T(p_k) - T(p_1)$,
$S(p'_k) = S(p_k) - S_0$ $(1 \leqq k \leqq n)$ (S_0 はPの初期シーケンス番号)

【００６４】

３）　パケットデータＤの先頭からパケット x を取り出す。パケット x のヘッダの (source IP, source port, destination IP, destination port)の組み合わせを見ることで、xの属するコネクションの一方の方向が分かる。x がその方向のm番目のパケットだとして、そのパケット系列を $X = \{ x_1, x_2, \ldots, x_m, \ldots \}$ と表す。（$x = x_m$ である。）

【００６５】

４）　l=1, 2, ..., m のそれぞれについて、以下の４－１），４－２），４－３）、４－４）を実行する。

【００６６】

４－１）　パケット系列 $X(m, l) = \{ x_l, x_{\{l+1\}}, \ldots, x_m \}$ のグラフの始点を原点にそろえたグラフをX'(m, l)とする。

$X'(m, l) = \{ x'_l, x'_{\{l+1\}}, \ldots, x'_m \}$,
$T(x'_k) = T(x_k) - T(x_l)$,
$S(x'_k) = S(x_k) - S(x_{\{l-1\}})$　$(l \leqq k \leqq m)$

【００６７】

４－２）　系列P'(n)のインデックスk(m)を、

$k(m) = \max\{ k \mid S(p'_k) \leqq S(x'_m) \}$

として、系列 X'(m, l)のグラフと系列P'(k(m))のグラフとが挟む領域（すなわち。この２つのグラフと横軸に平行な直線 $y=S(p'_{\{k(m)\}})$によって囲まれる領域）の面積 M(m, l) を求める。ただし、X'(m, l) のグラフがP'(k(m))のグラフの右側にある部分は正の面積、左側にある部分は負の面積として計算する。

【００６８】

【数１】

$$M(1, 1) = 0,$$
$$S(p'_0) = 0,$$
$$S(x'_0) = 0$$

$$M(m, 1) - M(m-1, 1)$$
$$= (T(x'_{\{m-1\}}) - T(p'_{\{k(m-1)+1\}})) \times (S(x'_{\{m-1\}}) S(p'_{\{k(m-1)\}}))$$
$$+ (T(x'_m) - T(p'_{\{k(m-1)+1\}})) \times (S(p'_{\{k(m-1\}+1\}}) - S(x'_{\{m-1\}}))$$
$$+ \Sigma_{(i=k(m-1)+2, \ldots, k(m))}((T(x'_m) - T(p'_i)) \times (S(p'_i) - S(p'_{\{i-1\}})))$$

【００６９】

４－３）　X'(m, 1) のグラフがP'(n)のグラフより最も右側にはみ出たときのタイムスタンプの差を U （正の値）、最も左側にはみ出たときのタイムスタンプの差を V （負の値）　として記憶しておく。上記４－２のM(m, 1) の時刻の差の計算において、

$T(x'_{\{j-1\}}) - T(p'_{\{k(j-1)+1\}})$,
$T(x'_j) - T(p'_{\{k(j-1)+1\}})$,
$T(x'_j) - T(p'_i)$　($k(j-1)+2 \leqq i \leqq k(j)$)
ただし, $j=1,2,\ldots,m$

のうち最も大きい値がU、最も小さい値がVである。

【００７０】

４－４）　X'(m, 1) のグラフの高さがP'(n)のグラフの高さ以上になったかどうか

$$k(m) \geqq n$$

で判定し、成立しなければ1 を一つ進めてステップ４－１へ戻る。

【００７１】

成立する場合はもうmをこれ以上大きくしても挟まれる領域は変わらないので、２つのグラフ X'(m, 1) , P'(n)が交差しないように X'(m, 1) を横方向へ移動して、移動後に２つのグラフによって挟まれる領域の面積を計算する。左方向へのUの移動と右方向へのVの移動の２つの場合のうち、小さいほうの面積を M(1

）にする。すなわち、

$$M(l) = \min \{ | M(m, l) - U \times S(p'_n) |, | M(m, l) - V \times S(p'_n) | \}$$

である．。この M(l) が M(0), M(1), ... M(l) の中で最小だったならばそれを改めて最小面積Mとして記録する。そのときのパケットの始まりのインデックスl、タイムスタンプの差（U又はV）を記憶する。１ を一つ進めてステップ４－１へ戻る。

【００７２】

５）　ステップ３、４をＤの中のパケットが無くなるまで繰り返して、全てのコネクションの中の各方向の系列Xについて系列Pとの最小面積Mから類似度 M/S(p'_n)をそれぞれ求める。これらを小さい順にソートして、類似度の高い順にコネクションの方向が得られる。

【００７３】

類似コネクションの特定及び返信.

各サイトのログ・ボックスは、上記の計算から得られた類似度が、予め定められた数値以下の系列をいくつか選択する。この選択された系列、この系列が含まれるコネクション、及びその類似度等の情報を、要求があったサイトに返信する。各サイトからのデータを受信した要求サイトは、それらの中から特に類似度が高いコネクションを見つける。それらは同一チェーン上のコネクションである可能性が高い。さらに、それらのホストの管理者と連絡を取り合い、本当にアタッカーに踏み台として使われたかどうかを確かめて、最終的には人手による追跡を行う。尚、次にアタッカーが侵入してくるときのことを考えて、類似度の高いコネクションの４つ組の中のＩＰアドレスが属するネットワークのバックボーンにも、パケットデータ記録用のＰＣを設置することにより、その後の監視を強化することができる。

【００７４】

尚、本実施形態においては、ログ・ボックスがパケット・データの記録及び、コネクションの比較判定の双方を行い、他のネットワーク上のコンピュータと通

信を行っている。しかし、通信内容を記録するログ・ボックスは、自分からは全く送信を行わない完全に受動的な機能のみに限定することにより、他者の侵入を防ぐことが可能となる。又、本実施形態は、不正侵入者を発見するために、アクセス・チェーンの追跡を行ったが、本発明は必ずしもこの用途に限定されるものではない。例えば、故障・欠陥があるコンピュータを探索するために、本実施形態の方法を使用することも可能である。

【図面の簡単な説明】

【図１】　従来における不正アクセスの構造を示す概略図である。

【図２】　実施の形態におけるネットワーク・システムを示す該略図である。

【図３】　実施の形態における不正アクセスの構造を示す概略図である。

【図４】　実施の形態におけるログ・ボックスのハードウェア構成を示す概略図である。

【図５】　実施の形態におけるログ・ボックスのソフトウェア構成を示す概略図である。

【図６】　実施の形態におけるパケットのデータ構造を示す概略図である。

【図７】　実施の形態におけるパケットのデータ構造を示す概略図である。

【図８】　実施の形態におけるパケット・データの構造を示す概略図である。

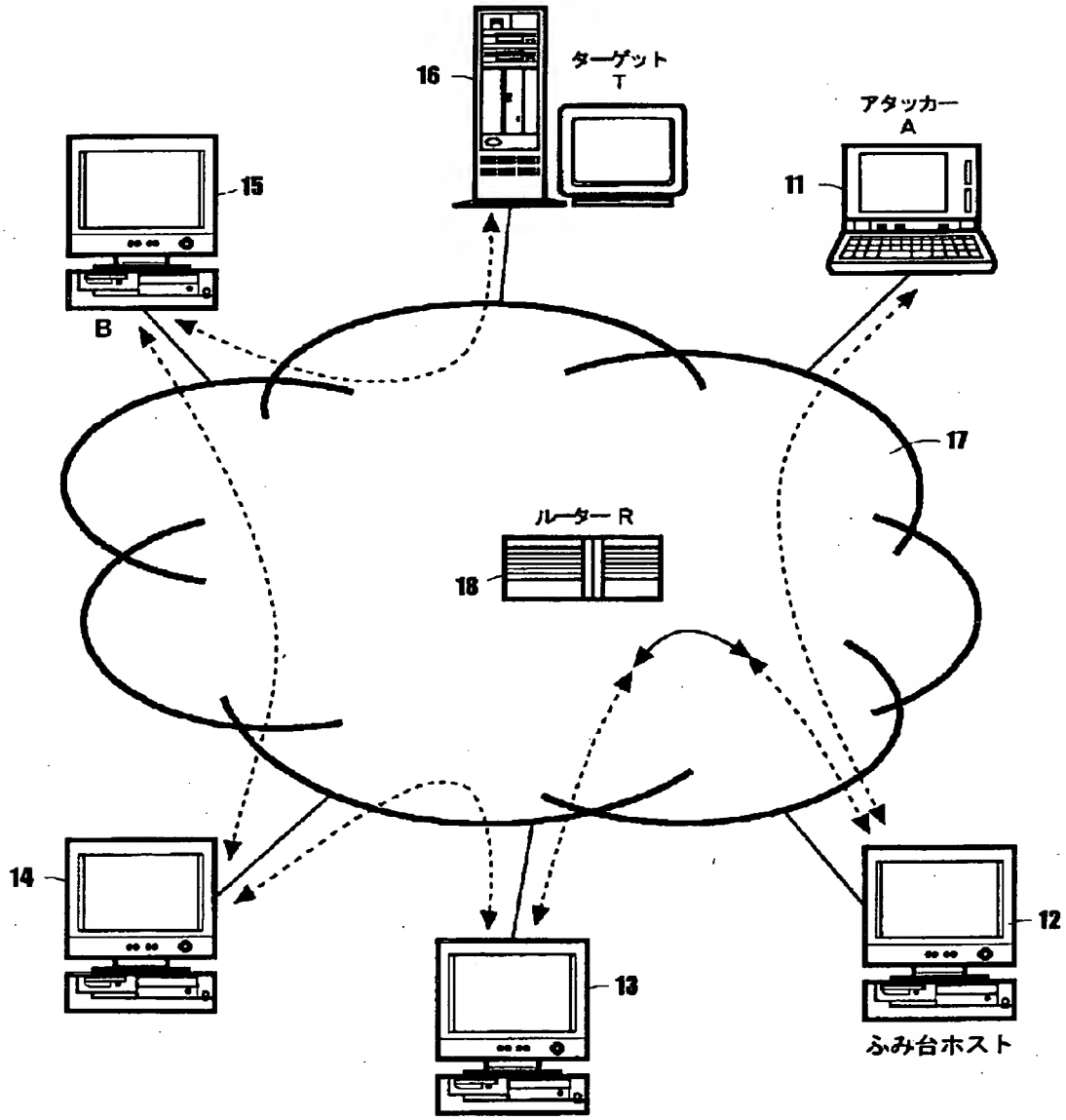
【図９】　実施の形態におけるを示す比較判定部のソフトウェア構成を示す概略図である。

【図１０】　実施の形態におけるを示す系列比較方法を説明するためのグラフである。

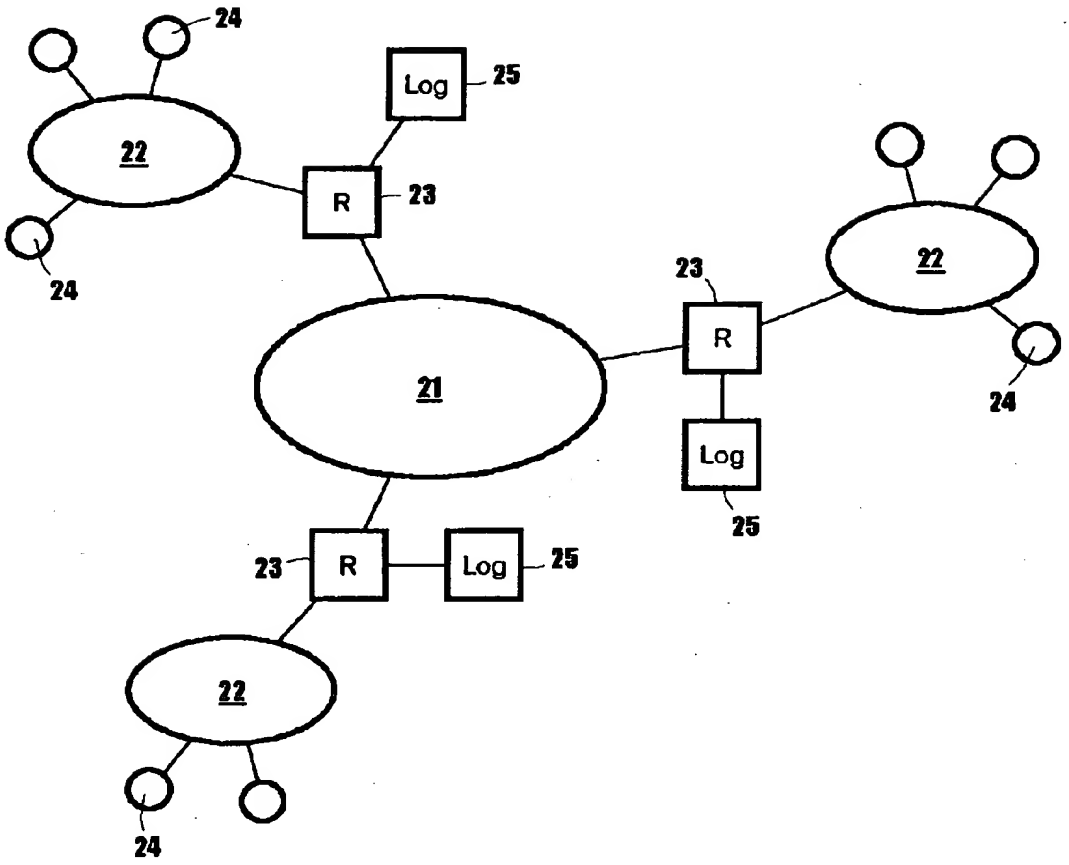
【符号の説明】　１１　アタッカー、１２、１３、１４、１５　踏み台ホスト、１６　ターゲット、１７　ネットワーク、１８　ルータ、２１　バックボーン、２２　ネットワーク、　２３　ルータ、２４　端末、２５　ログボックス、３１　アタッカー、３２、３３、３４、３５　踏み台ホスト、３６　ターゲット、３７、３８、３９　ログボックス、４０　ルータ

【書類名】　　図面

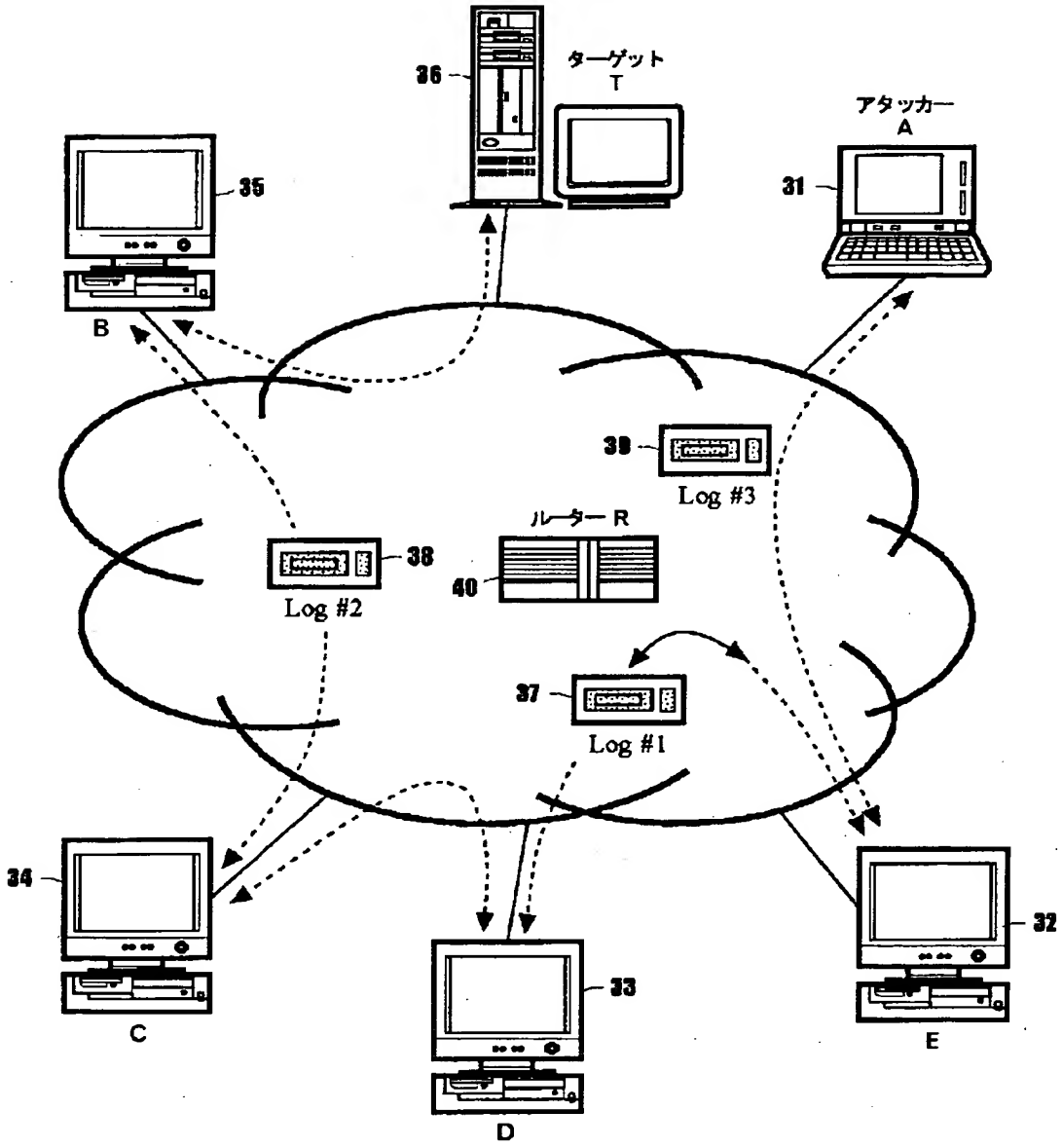
【図１】



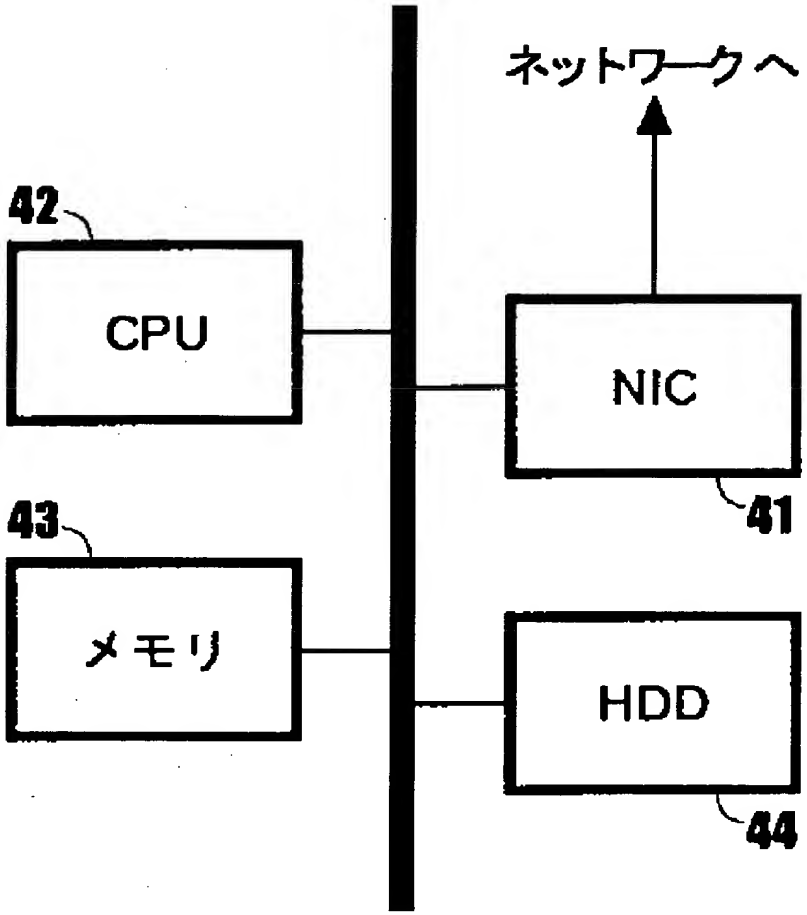
【図２】



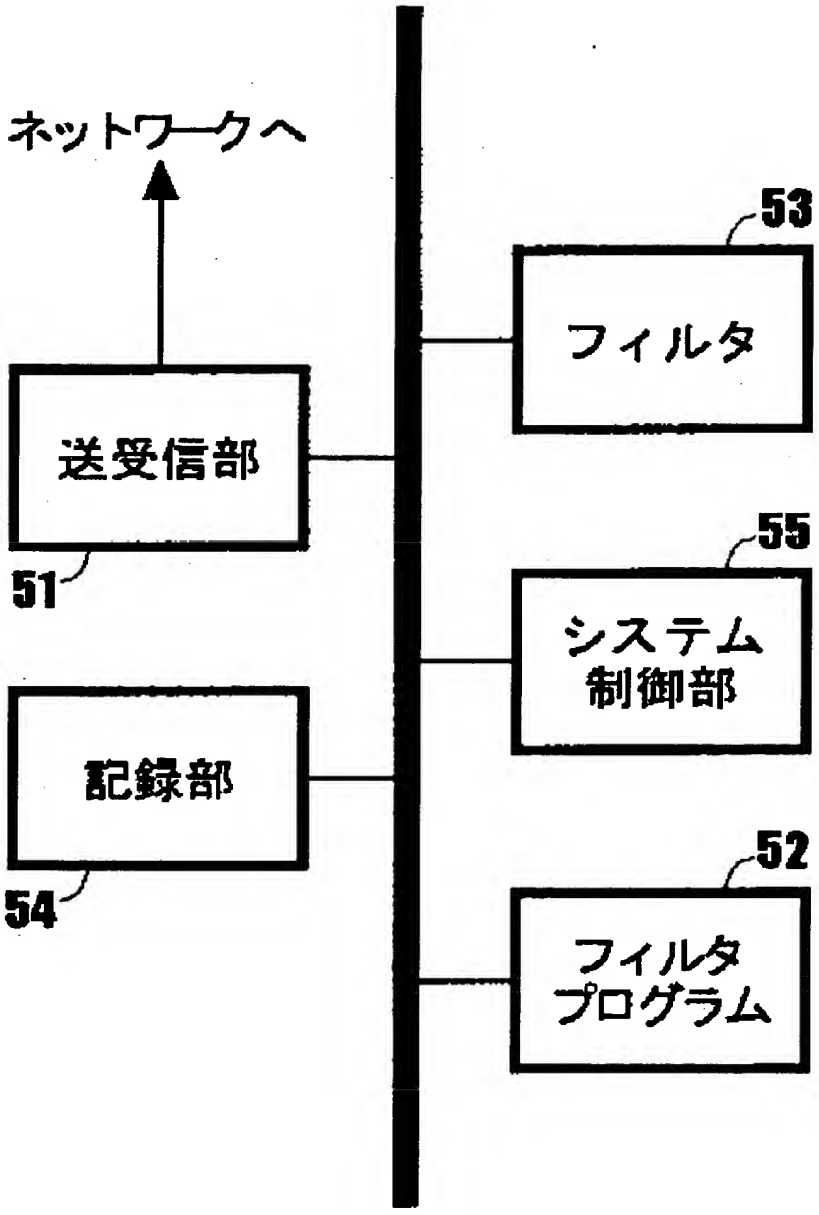
【図３】


36
ターゲット
T
アタッカー
A
31
35
B
39
Log #3
ルーター R
38
Log #2
40
37
Log #1
34
C
33
D
32
E

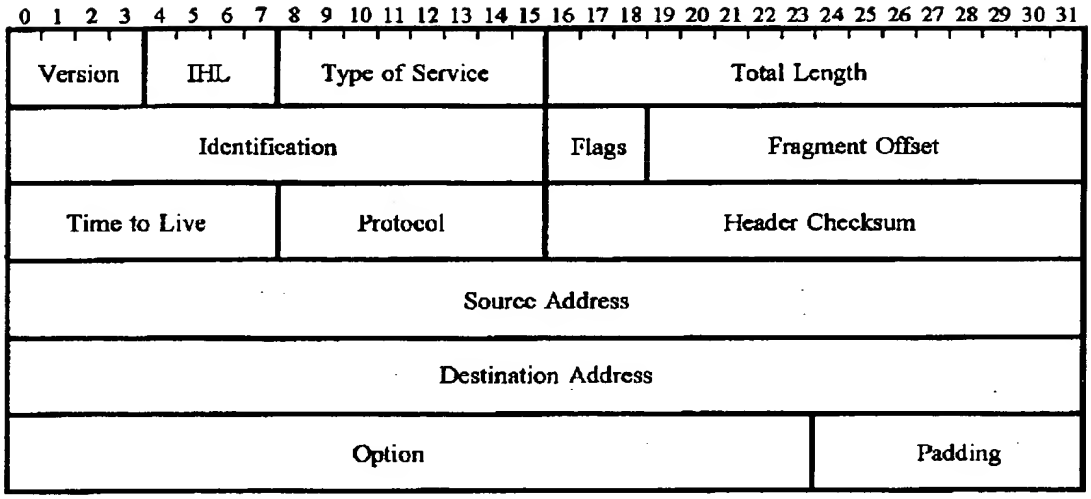
【図４】



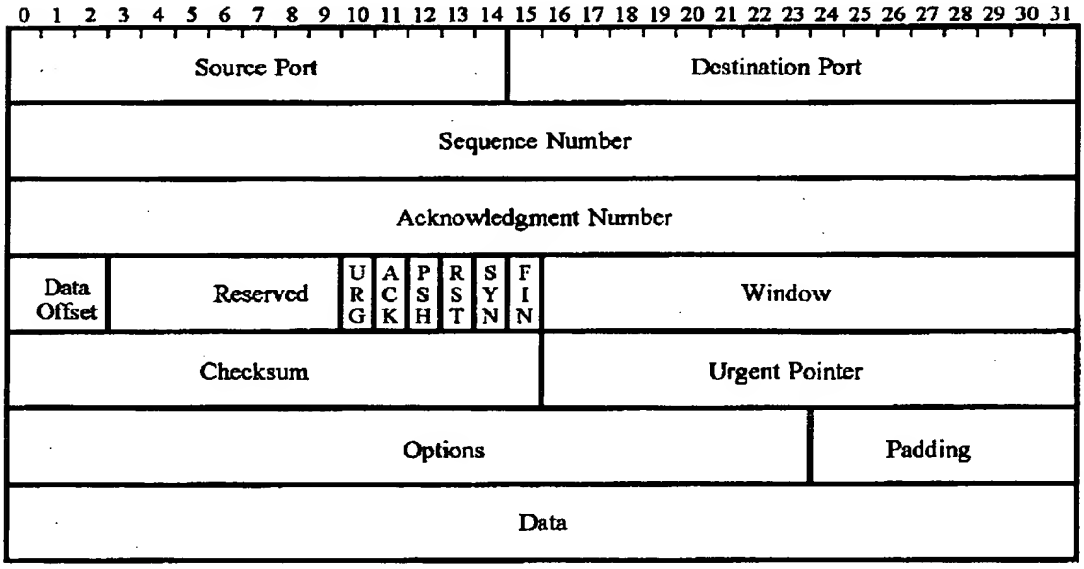
【図５】



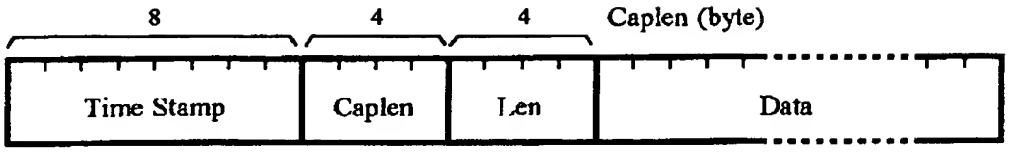
【図６】


0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Version
IHL
Type of Service
Total Length
Identification
Flags
Fragment Offset
Time to Live
Protocol
Header Checksum
Source Address
Destination Address
Option
Padding

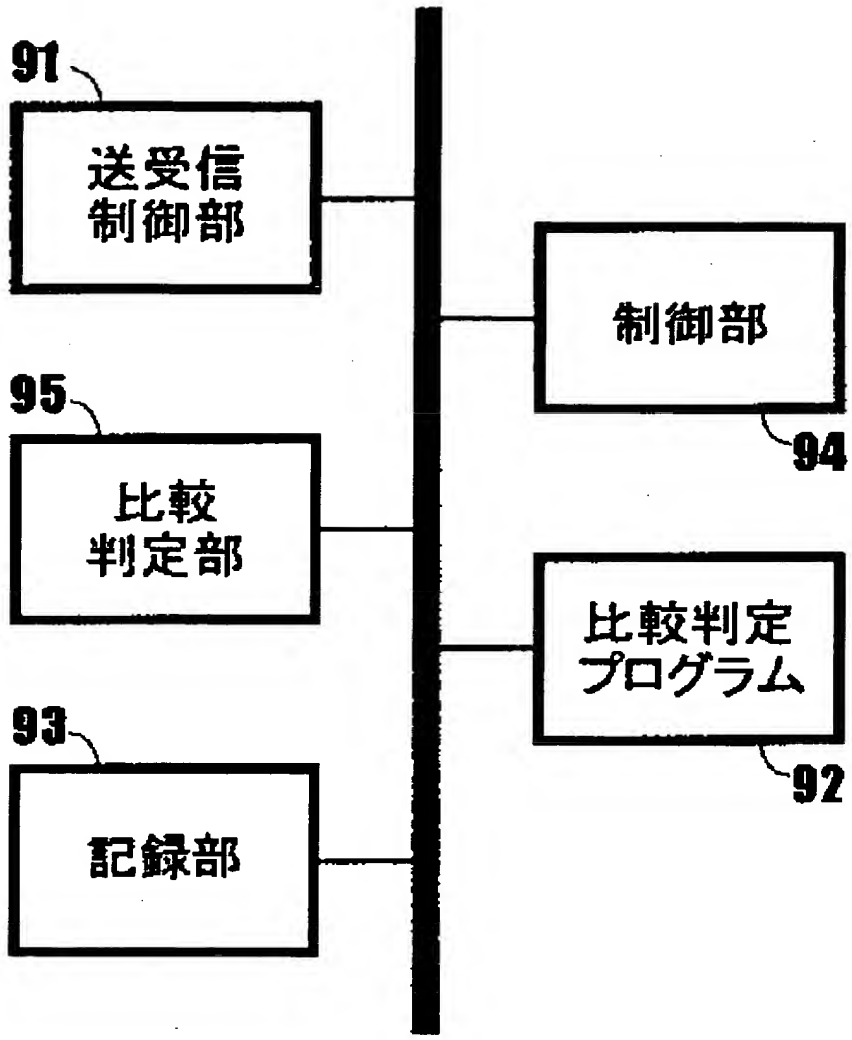
【図７】


0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Source Port
Destination Port
Sequence Number
Acknowledgment Number
Data Offset
Reserved
URG
ACK
PSH
RST
SYN
FIN
Window
Checksum
Urgent Pointer
Options
Padding
Data

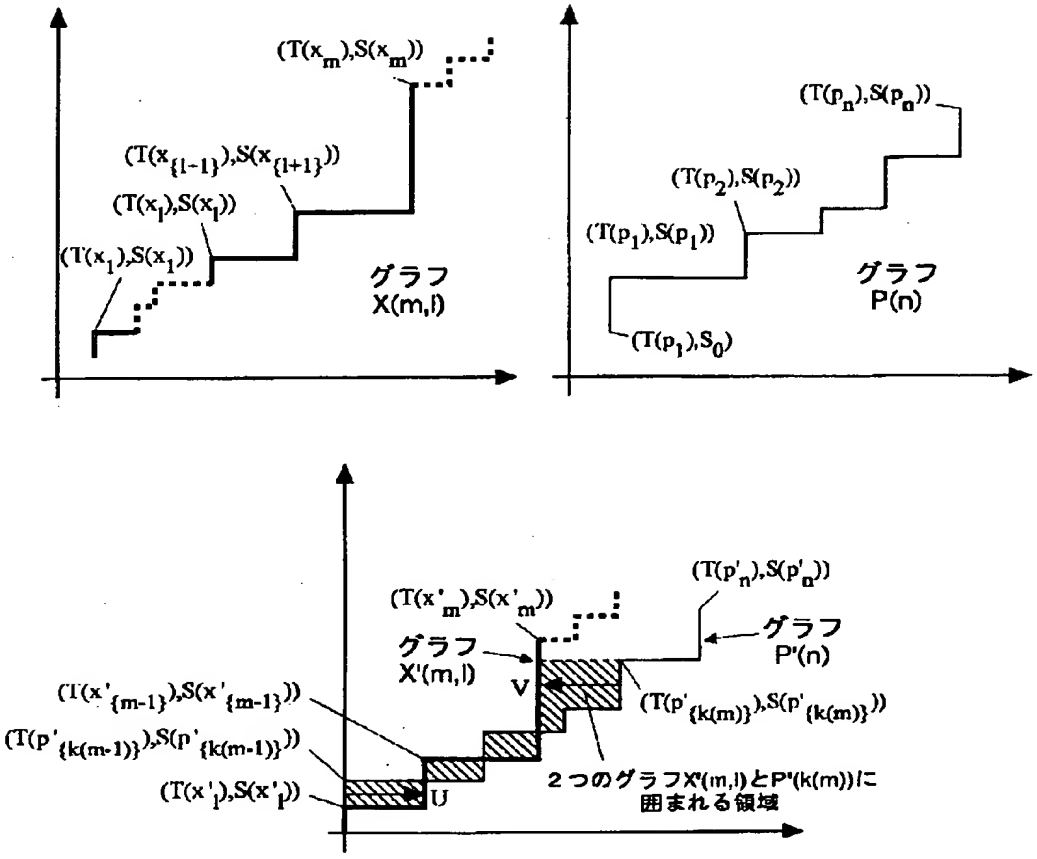
【図８】



【図９】



【図１０】


(T(x_m),S(x_m))
(T(x_{l-1}),S(x_{l+1}))
(T(x_l),S(x_l))
(T(x_1),S(x_1))
グラフ
X(m,l)
(T(p_n),S(p_n))
(T(p_2),S(p_2))
(T(p_1),S(p_1))
(T(p_1),S_0)
グラフ
P(n)
(T(x'_m),S(x'_m))
(T(p'_n),S(p'_n))
グラフ
X'(m,l)
グラフ
P'(n)
V
(T(x'_{m-1}),S(x'_{m-1}))
(T(p'_{k(m)}),S(p'_{k(m)}))
(T(p'_{k(m-1)}),S(p'_{k(m-1)}))
(T(x'_1),S(x'_1))
U
２つのグラフX'(m,l)とP'(k(m))に
囲まれる領域

【書類名】　　要約書

【要約】

【課題】　パケットのデータ内容に依存しない、アクセス・チェーンの追跡方法及びそのシステムを得る。

【解決手段】　ネットワークを行き交うパケットのログデータを、ログボックスに記録する。この時、パケットのデータサイズと検知時刻を記録する。ターゲット・コンピュータに対して不正アクセスがあった場合、このログ情報に基づいて、不正アクセスチェーンを追跡する。このアクセスチェーンの追跡は、以下のように行う。ログデータから、第１のコネクションにおける時刻によるパケットのデータ・サイズの変化と、前記第２のコネクションにおける時刻によるパケットのデータ・サイズの変化とを算出する。これらの系列によって形成されるグラフの形状を比較し、類似してる場合は同一のチェーンに含まれるコネクションであると判定する。

【選択図】　　図２

認定・付加情報

特許出願の番号	特願２０００－０２５５９４
受付番号	５０００００１１６２４１
書類名	特許願
担当官	第七担当上席　　００９６
作成日	平成１２年　２月　３日

＜認定情報・付加情報＞

【提出日】　平成12年 2月 2日

次頁無

出　願　人　履　歴　情　報

識別番号　　〔３９０００９５３１〕

１．変更年月日　　１９９０年１０月２４日
［変更理由］　新規登録
住　所　　アメリカ合衆国１０５０４、ニューヨーク州　アーモンク　（番地なし）
氏　名　　インターナショナル・ビジネス・マシーンズ・コーポレイション

２．変更年月日　　２０００年　５月１６日
［変更理由］　名称変更
住　所　　アメリカ合衆国１０５０４、ニューヨーク州　アーモンク　（番地なし）
氏　名　　インターナショナル・ビジネス・マシーンズ・コーポレーション